Balanced collective contributions, the equal allocation of non-separable costs and application to data sharing games

Sylvain Béal, Marc Deschamps, Philippe Solal

Balanced collective contributions, the equal allocation of non-separable costs and application to data sharing games

Sylvain Béal, Marc Deschamps, Philippe Solal

February 2014

Working paper No. 2014 – 02

CRESE
30, avenue de l'Observatoire
25009 Besançon
France
http://crese.univ-fcomte.fr/

Balanced collective contributions, the equal allocation of non-separable costs and application to data sharing games *

Sylvain Béal ‡, Marc Deschamps §, Philippe Solal ¶

August 28, 2014

Abstract

The axiom of Balanced collective contributions is introduced as a collective variant of the axiom of Balanced contributions proposed by Myerson (1980). It requires that the identical average impact of the withdrawal of any agent from a game on the remaining population. It turns out that Balanced collective contributions and the classical axiom of Efficiency characterize the equal allocation of non-separable costs, an allocation rule which is extensively used in cost allocation problems and in accounting. For instance, the equal allocation of non-separable costs coincides with the Nucleolus on the class of data sharing games within the European REACH legislation. While our result does not hold on data sharing games, we provide comparable characterizations of the equal allocation of non-separable costs and the Shapley value.

Keywords: Balanced collective contributions, Balanced contributions, Equal allocation of non-separable costs, Shapley value, Data games.

JEL Classification number: C71, D71, K32, L65.

1 Introduction

Cost allocation problems arise in many real life situations, where individuals, all with their own purposes, decide to work together. In these situations the problem is to divide among the participants the joint costs which result from the cooperation. Such situations can be modeled by cooperative games with transferable utilities, for which a large number of allocation rules exist. The most famous such allocation rule is maybe the Shapley value (Shapley, 1953), which is based on the marginal contributions of the involved agents to the coalitions they belong. An attractive axiomatic characterization of the Shapley value by efficiency and balanced contributions is provided by Myerson (1980). Efficiency requires that the cost incurred by the coalition of all agents is fully divided among them. Balanced contributions requires, for any two agents, equal allocation variation after the leave of the other agent.

In this article, we introduce a collective variant of balanced contributions. Balanced collective contributions imposes the same requirement as balanced contribution, except that the allocation variation is averaged on the remaining population of agents. In other words, balanced collective contributions requires, for any two agents, equal average allocation variation in the population that remains after the leave of the other agent. As such, our axiom is silent on the exact individual impacts within these remaining populations. It possesses a similar flavor to the axiom of component fairness on the class of cooperative games with a cycle-free communication graph as defined by Herings et al. (2008), which requires average allocation variation for the two components of the graph created by the deletion of a link. Combined with efficiency, it turns out that balanced collective contributions characterizes another well-known allocation rule called the equal

*Financial support by the National Agency for Research (ANR) – research program "DynaMITE: Dynamic Matching and Interactions: Theory and Experiments", contract ANR-13-BSHS1-0010 – and the "Mathématiques de la décision pour l'ingénierie physique et sociale" (MODMAD) project is gratefully acknowledged.

‡Corresponding author. Université de Franche-Comté, CRESE, 30 Avenue de l'Observatoire, 25009 Besançon, France. E-mail: sylvain.beal@univ-fcomte.fr. Tel: (+33)(0)3-81-66-68-26. Fax: (+33)(0)3-81-66-65-76

§Université de Nice Sophia-Antipolis, GREDEG (CNRS UMR 7321) and BETA (CNRS 7522), France. E-mail: marc.deschamps@univ-lorraine.fr.

¶Université de Saint-Etienne, CNRS UMR 5824 GATE Lyon Saint-Etienne, France. E-mail: philippe.solal@univ-st-etienne.fr. Tel: (+33)(0)4-77-42-19-61. Fax: (+33)(0)4-77-42-19-50.

allocation of non-separable costs (Proposition 1). This allocation rule is based on the marginal contribution of the agents to the grand coalition, also called the separable costs. The non-separable costs is what remains of the cost incurred by the grand coalition after deleting the sum of these marginal contributions. The equal allocation of non-separable costs first assigns to each agent its separable cost and then split equally the non-separable costs.

The equal allocation of non-separable costs has received considerable attention is cost allocation problems. For instance, it has been successfully applied by the Tennessee Valley Authority (TVA) in the 1930's to sharing costs of dam systems along the Tennessee River (Ransmeier, 1942; Heaney, 1979; Straffin and Heaney, 1981). Furthermore, for the class of 1-concave games, Driessen (1988) has shown that the equal allocation of non-separable costs coincides with the Nucleolus (Schmeidler, 1969). As a special case, the class of 1-concave games contains the library games considered in Driessen et al. (2012) to study the Catalan university library consortium for subscription to journals issued by Kluwer publishing house, and the data games considered in Dehez and Tellone (2013) to study the determination of compensations associated with the data sharing problem in the European REACH legislation.

As an application, we consider the class of data games. Thomson (2001, page 343) emphasizes that

> "It is important to understand how a characterization is affected by enlarging or restricting the domain of problems under consideration."

On the smaller class of data games, we show that the uniqueness results of Proposition 1 and in Myerson (1980) do not hold. Moreover, balanced collective contributions and balanced contributions can be required only for those data games for which all involved subgames are data games too. Proposition 2 proves that adding two extra axioms proposed in Béal and Deschamps (2014) allows to recover a characterization of the equal allocation of non-separable costs on the class of data games. Surprisingly, the two added axioms help to characterize the Shapley value on the class of data games when balanced collective contributions is replaced by balanced contributions (Proposition 3).

The rest of the article is is organized as follows. Section 2 provides definitions and notations. The characterization of the equal allocation of non-separable costs on the class of all cost games is presented in section 3. Section 4 is devoted to the application to data games. Finally, section 5 concludes.

2 Definitions and notations

Let $\mathcal{U} \subseteq \mathbb{N}$ be a fixed and infinite universe of agents. Denote by U the set of all finite subsets of $\mathcal{U}$. A **cost game** is a pair (N, c) where $N \in U$ and $c : 2^N \longrightarrow \mathbb{R}$ such that $c(\emptyset) = 0$. A non-empty subset $S \subseteq N$ is a coalition, and $c(S)$ is the minimal costs which should be involved if the individuals in S should work together in order to serve their own purposes. For any non-empty coalition S, let s be the cardinality of S. The **sub-game** of (N, c) induced by $S \subseteq N$ is denoted by $(S, c_{|S})$ and define as $c_{|S}(T) = c(S)$ for all $T \in 2^S$. Define $\mathcal{C}$ as the class of all cost games with a finite agent set in $\mathcal{U}$. An **allocation rule** on $\mathcal{C}$ is a function f that assigns a payoff vector $f(N, c) \in \mathbb{R}^N$ to any $(N, c) \in \mathcal{C}$. In this article, we consider the following allocation rules.

The **Shapley value** (Shapley, 1953) is the allocation rule Sh defined as:

$$\forall (N, c) \in \mathcal{C}, \forall i \in N, \quad \mathrm{Sh}_i(N, c) = \sum_{S \in 2^N : S \ni i} \frac{(s-1)!(n-s)!}{n!} \big(c(S) - c(S \backslash \{i\})\big).$$

Agent i's separable cost is his marginal contribution to the grand coalition $c(N) - c(N \backslash \{i\})$, so that the non-separable costs are

$$c(N) - \sum_{j \in N} \big(c(N) - c(N \backslash \{j\})\big).$$

The equal allocation of non-separable costs first assigns to each agent his separable costs, and then split equally the non-separable costs among the agents. Formally, the **Equal Allocation of Non-separable Costs** is the allocation rule EANSC defined as:

$$\forall (N,c) \in \mathcal{C}, \forall i \in N, \quad \text{EANSC}_i(N,c) = c(N) - c(N\backslash\{i\}) + \frac{1}{n}\left(c(N) - \sum_{j \in N}\big(c(N) - c(N\backslash\{j\})\big)\right). \qquad (1)$$

3 Axiomatic characterization

In this section we invoke the following axioms. The first two axioms are classical.

Efficiency. For all $(N,c) \in \mathcal{C}$, $\sum_{i \in N} f_i(N,c) = c(N)$.

Balanced contributions. For all $(N,c) \in \mathcal{C}$, all $i,j \in N$,

$$f_j(N,c) - f_j(N\backslash\{i\}, c_{|N\backslash\{i\}}) = f_i(N,c) - f_i(N\backslash\{j\}, c_{|N\backslash\{j\}}).$$

Balanced contributions requires that if a agent leaves a game, then the payoff variation for another agent is identical to his own payoff variation if this other agent leaves the game. Myerson (1980) characterizes the Shapley value on $\mathcal{C}$ by Efficiency and Balanced Contributions. The third axiom, called Balanced collective contributions, is inspired by Myerson (1980)'s Balanced contributions. Balanced collective contributions is build on the same principle, except that the payoff variation is measured collectively, by averaging on all remaining agents.

Balanced collective contributions. For all $(N,c) \in \mathcal{C}$, all $i,j \in N$,

$$\frac{1}{n-1}\sum_{k \in N\backslash\{i\}}\Big(f_k(N,c) - f_k(N\backslash\{i\}, c_{|N\backslash\{i\}})\Big) = \frac{1}{n-1}\sum_{k \in N\backslash\{j\}}\Big(f_k(N,c) - f_k(N\backslash\{j\}, c_{|N\backslash\{j\}})\Big).$$

It turns out that the equal allocation of non-separable costs is characterized by Efficiency and Balanced collective contributions.

Proposition 1 *The equal allocation of non-separable costs is the unique allocation rule on $\mathcal{C}$ that satisfies Efficiency and Balanced collective contributions.*

Proof. It is obvious that the EANSC satisfies Efficiency. Regarding Balanced collective contributions, choose any $(N,c) \in \mathcal{C}$, and any $i \in N$. By Efficiency in the subgame $(N\backslash\{i\}, c_{|N\backslash\{i\}})$, we have

$$\sum_{k \in N\backslash\{i\}} \text{EANSC}_k(N\backslash\{i\}, c_{|N\backslash\{i\}}) = c_{|N\backslash\{i\}}(N\backslash\{i\}) = c(N\backslash\{i\}),$$

so that

$$\frac{1}{n-1}\sum_{k \in N\backslash\{i\}}\Big(\text{EANSC}_k(N,c) - \text{EANSC}_k(N\backslash\{i\}, c_{|N\backslash\{i\}})\Big) = \frac{1}{n-1}\left(\sum_{k \in N\backslash\{i\}} \text{EANSC}_k(N,c) - c(N\backslash\{i\})\right).$$

By another application of Efficiency in (N,c) the previous expression becomes

$$\frac{1}{n-1}\left(c(N) - \text{EANSC}_i(N,c) - c(N\backslash\{i\})\right),$$

or equivalently, by definition (1) of EANSC,

$$\frac{1}{n-1}\frac{1}{n}\left(c(N)-\sum_{j\in N}\big(c(N)-c(N\backslash\{j\})\big)\right).$$

Since this expression does not depend on the chosen agent $i \in N$, this proves that EANSC satisfies Balanced collective contributions.

It remains to show that if an allocation rule f on $\mathcal{C}$ satisfies Efficiency and Balanced collective contributions, then it coincides with EANSC. So, consider such an allocation rule and pick any game $(N, c) \in \mathcal{C}$. If $n = 1$, then the assertion follows from Efficiency. So assume that $n \geq 2$ and consider any pair of agents $i, j \in N$. By Balanced collective contributions, it holds that

$$\frac{1}{n-1}\sum_{k\in N\backslash\{i\}}\Big(f_k(N,c)-f_k(N\backslash\{i\},c_{|N\backslash\{i\}})\Big)=\frac{1}{n-1}\sum_{k\in N\backslash\{j\}}\Big(f_k(N,c)-f_k(N\backslash\{j\},c_{|N\backslash\{j\}})\Big).$$

By Efficiency of f in (N, c), $(N\backslash\{i\}, c_{|N\backslash\{i\}})$ and $(N\backslash\{j\}, c_{|N\backslash\{j\}})$, and the definition of a subgame, we can rewrite the previous equality as

$$c(N)-f_i(N,v)-c(N\backslash\{i\})=c(N)-f_j(N,v)-c(N\backslash\{j\}),$$

or equivalently

$$f_i(N,c)-\big(c(N)-c(N\backslash\{i\})\big)=f_j(N,c)-\big(c(N)-c(N\backslash\{j\})\big).$$

Summing over all $j \in N$, we get

$$n\big(f_i(N,c)-\big(c(N)-c(N\backslash\{i\})\big)\big)=\sum_{j\in N}\big(f_j(N,c)-\big(c(N)-c(N\backslash\{j\})\big)\big),$$

and by Efficiency of f in (N, c),

$$n\big(f_i(N,c)-\big(c(N)-c(N\backslash\{i\})\big)\big)=c(N)-\sum_{j\in N}\big(c(N)-c(N\backslash\{j\})\big).$$

Rearranging, we obtain

$$f_i(N,c)=c(N)-c(N\backslash\{i\})+\frac{1}{n}\left(c(N)-\sum_{j\in N}\big(c(N)-c(N\backslash\{j\})\big)\right)=\mathrm{EANSC}_i(N,c).$$

Since $(N, c) \in \mathcal{C}$ and $i \in N$ were arbitrary chosen, the proof is complete. ■

The two axioms invoked in Proposition 1 are logically independent since the Shapley value satisfies Efficiency but not Balanced collective contributions while the null solution, which assigns a null payoff vector to all games satisfies Balanced collective contributions but not Efficiency.

4 Application to data sharing and the REACH legislation

The European institutions have decided, since December 13, 2006, the implementation of a new harmonized legislative framework in the field of chemical industry: the European Regulation (EC) No 1907/2006 REACH (Registration, Evaluation, Authorization and Restriction of Chemicals), which ultimately aims to ensure greater human and environmental safety (based on impact studies by the European Commission it will reduce the number of deaths due to cancer by a range of 2000 to 4000 per year and lead to a reduction in public health spending of up to 50 billion over thirty years), while preserving and enhancing the competitiveness of the European chemical industry.

Article 30 of Regulation (EC) No 1907/2006 concerns the sharing of data between users of a chemical substance. The most innovative aspect of this Regulation lies in the fact that companies are forced to provide all existing data on the properties of the chemicals they use. According to the procedure, without data companies cannot use the substances. The Regulation is based on the "no data, no market rule", *i.e.* chemical substances may in principle not be manufactured in the EU or placed on the market unless they have been registered. Any company wishing to declare a substance must participate in a SIEF (Substance Information Exchange Forum). Within a SIEF, members are free to decide on legal form, organization and communication modalities. The Regulation insists on the fact that registrants "shall make every effort to ensure that the costs of sharing the information are determined in a fair, transparent and non-discriminatory way" (Article 27, §3). The question of how to realize data sharing within a SIEF, and to determine the necessary compensations is therefore a critical issue.

In this section, we study this question by means of a special class of cost games called data games (Dehez and Tellone, 2013).[1] For such problems, compensation schemes specify how the data owners should be compensated by the agents in needs of data. Since data games are 1-concave games in the sense of Driessen (1988) (see also Driessen and Khmelnitskaya, 2013), the equal allocation of non-separable costs coincides with the Nucleolus (see Schmeidler, 1969). Therefore, Propositions 2 and 3 provide alternative comparable characterizations of the equal allocation of non-separable costs and the Shapley value to those in Béal and Deschamps (2014). Apart from differences in the axiom sets, our results differ from those in Béal and Deschamps (2014) in that they are obtained on the class of data games with variable agent sets, while the agent set is fixed in Béal and Deschamps (2014).

Formally, let $\mathcal{D}$ be a nonempty finite universe of data. Each data $k \in \mathcal{D}$ is characterized by a cost $d_k \in \mathbb{R}_+$, which is interpreted as the cost of replicating the data. A **data sharing problem** is a described by a pair (N, D), where

- $N \in U$ represents the agents involved in a SIEF.
- $D \subseteq \mathcal{D}$ is a nonempty finite set of data. Data in D are relevant to the study of the chemical substance associated with the SIEF. Subscripts k and h are used to refer to data. Denote by D_i the data owned by agent $i \in N$ in D. It is assumed that each data in D is held by at least one agent in N.

We denote by $\mathcal{P}$ the set of all data sharing problems with a finite agent set. For each nonempty coalition of agents $S \in 2^N$, $D_S = \cup_{i\in S} D_i$ stands for the set of all data held by the members of S. We shall keep the notations D_i instead of $D_{\{i\}}$ and D instead of D_N. For each $i \in N$, D_i^E is the set of data in D that i exclusively holds, *i.e.* the set, possibly empty, of data held by i and by no other agent in $N\backslash\{i\}$. Thus $D_i^E = D\backslash D_{N\backslash\{i\}}$. Define $D^E = \cup_{i\in N} D_i^E$ as the set of exclusive data in D. Finally, denote by $o_k(N, D)$ the number of owners of data $k \in D$ among the agent set N, *i.e.* $o_k(N, D) = |\{i \in N : k \in D_i\}|$.

To each data sharing problem $(N, D) \in \mathcal{P}$, it is useful to associate a **data game** $(N, c_D) \in \mathcal{C}$, where the characteristic function $c_D : 2^N \longrightarrow \mathbb{R}_+$ assigns to each nonempty coalition $S \in 2^N$ the total cost $c_D(S) \in \mathbb{R}_+$ of the data in D that the members of S do not hold. Formally, for each $S \in 2^N$, $S \neq \emptyset$:

$$c_D(S) = \sum_{k\in D\backslash D_S} d_k,$$

and by convention $c_D(\emptyset) = 0$. Observe that $c_D(N) = 0$ since every data in D is held by some agent in N. We denote by $\mathcal{DC}$ the set of all data games that can be constructed from $\mathcal{P}$, *i.e.* $\mathcal{DC} = \{(N, c_D) \in \mathcal{C} : (N, D) \in \mathcal{P}\}$.

[1] For other cooperative approaches to the problem of collecting information, we refer the reader to Brânzei et al. (2006) and the references therein.

Thus $\mathcal{DC} \subsetneq \mathcal{C}$, *i.e.* the class of data games is a strict subset of the class of cost games. Each data $k \in \mathcal{D}$ generates **elementary data games** $(N, c_{\{k\}}) \in \mathcal{DC}$.

For a data game $(N, c_D) \in \mathcal{DC}$, if an allocation rule assigns a positive payoff to an agent, then this agent has to pay a compensation, and otherwise he/she receives a compensation. On the class of data games, the equal allocation of non-separable costs can be rewritten as:[2]

$$\forall (N, c_D) \in \mathcal{DC}, \forall i \in N, \quad \text{EANSC}_i(N, c_D) = \sum_{k \in D^E} \frac{d_k}{n} - \sum_{k \in D_i^E} d_k. \tag{2}$$

The Shapley value admits a similar formulation:

$$\forall (N, c_D) \in \mathcal{DC}, \forall i \in N, \quad \text{Sh}_i(N, c_D) = \sum_{k \in D} \frac{d_k}{n} - \sum_{k \in D_i} \frac{d_k}{o_k(N, D)}. \tag{3}$$

The Shapley value contains two parts. In the first part, each agent contributes to a communal fund through the same fraction $1/n$ of the total cost of data. In the second part, the accumulated fund is redistributed data by data at an equal rate to each owner. It is easy to figure out that the equal allocation of non-separable costs implements the same principle, but only for the exclusive data.

In order to invoke the axiom of Balanced collective contributions on the class of data games, it is necessary to discuss the notion of data subgames. For any $(N, c_D) \in \mathcal{DC}$ and any nonempty $S \in 2^N$, by definition of the subgame $(S, c_{D|S})$ induced by S, we have $c_{D|S}(T) = \sum_{k \in D \setminus D_T} d_k$ for all nonempty $T \in 2^S$. Therefore, the subgame $(S, c_{D|S})$ induced by S is constructed from the pair (S, D): the data set under consideration is the same, but the agents in $N \setminus S$ are not involved anymore. By assumption, the pair (S, D) is a data sharing problem only if every data in D belongs to some agent in S, *i.e.* if the agents in $N \setminus S$ do not own exclusive data. As a consequence, the subgame $(S, c_{D|S})$ of (N, D) induced by S is a data game only if $D^E_{N \setminus S} = \emptyset$, or equivalently $D_S = D$. Otherwise, we would have $c_{D|S}(S) > 0$ so that $(S, c_{D|S})$ would not belong to $\mathcal{DC}$. Note also that whenever $D^E_{N \setminus S} = \emptyset$, it is the case that $(S, c_{D|S}) = (S, c_D)$. Turning to the subgames appearing in Balanced collective contributions, for a data game $(N, c_D) \in \mathcal{DC}$ and an agent $i \in N$, the subgame $(N \setminus \{i\}, c_{D|N \setminus \{i\}})$ induced by $N \setminus \{i\}$ is therefore a data game in $\mathcal{DC}$ if and only if $D_i^E = \emptyset$. In other words, the axiom of Balanced collective contributions cannot be applied to all data games in $\mathcal{DC}$, but only to those data games in which no data is exclusively held. From all these remarks, an equivalently statement of the axiom on $\mathcal{DC}$ can be formulated as follows:

Balanced collective contributions. For all $(N, c_D) \in \mathcal{DC}$ such that $D^E = \emptyset$, all $i, j \in N$,

$$\frac{1}{n-1} \sum_{k \in N \setminus \{i\}} \Big(f_k(N, c_D) - f_k(N \setminus \{i\}, c_D) \Big) = \frac{1}{n-1} \sum_{k \in N \setminus \{j\}} \Big(f_k(N, c_D) - f_k(N \setminus \{j\}, c_D) \Big).$$

Another evident consequence of all these observations is that Proposition 1 does not hold on the class $\mathcal{DC}$. In fact, the equal allocation of non-separable costs is not the unique allocation rule on $\mathcal{DC}$ that satisfies Efficiency and Balanced collective contributions since the null solution also satisfies these two axioms on $\mathcal{DC}$. In order to recover a characterization of the equal allocation of non-separable costs it is necessary to invoke extra axioms. It turns out that two such axioms already proposed in Béal and Deschamps (2014) can be used. We need the following notations. For a data game $(N, c_D) \in \mathcal{C}$ and an agent $i \in N$, we denote by

$$b_i^{\max}(N, c_D) = \max_{S \in 2^N : i \in S} \Big(c_D(S) - c_D(S \setminus \{i\}) \Big)$$

[2] Since an empty sum as a zero value by convention, observe that $\text{EANSC}(N, c_D) = \mathbf{0}_n$ whenever $D^E = \emptyset$.

and

$$b_i^{\min}(N, c_D) = \min_{S \in 2^N : i \in S} \Big(c_D(S) - c_D(S \backslash \{i\}) \Big)$$

agent i's maximal (or worst) and minimal (or best) marginal contributions to the coalitions he belongs to, respectively.

Equal concessions. $\forall (N, c_D) \in \mathcal{C}$ such that $D^E = D$, $\forall i, j \in N$, $f_i(N, c_D) - b_i^{\min}(N, c_D) = f_j(N, c_D) - b_j^{\min}(N, c_D)$.

Pooling. $\forall (N, c_D) \in \mathcal{C}$ and D', D'' such that $D' \cup D'' = D$ and $D' \cap D'' = \emptyset$, $f(N, c_D) = f(N, c_{D'}) + f(N, c_{D''})$.

Equal concessions only deals with data games in which every data is exclusive to some agent. The best contribution of an agent to the coalitions he can belong may be interpreted as his so-called utopia claim (especially because most of time the sum of the corresponding compensations is unrealistic, *i.e.* much less than zero). With this view in mind, equal concessions imposes that all agents abandon exactly the same amount from their utopia claims. Pooling is simply an adaptation of the classical axiom of additivity to data games. The result below shows that the equal allocation of non-separable costs is the unique allocation rule on the class of data games that satisfies Efficiency[3], Balanced collective contributions, Pooling and Equal Concessions.

Proposition 2 *The equal allocation of non-separable costs is the unique allocation rule on $\mathcal{DC}$ that satisfies Efficiency, Balanced collective contributions, Pooling and Equal concessions.*

Proof. It is obvious that EANSC satisfies Efficiency and Pooling, and it follows from Proposition 1 that EANSC satisfies Balanced collective contributions on $\mathcal{DC}$. Regarding Equal concessions, consider any $(N, c_D) \in \mathcal{DC}$ such that $D^E = D$. By (2), this means that

$$\forall i \in N, \quad \text{EANSC}_i(N, c_D) = \sum_{k \in D} \frac{d_k}{n} - \sum_{k \in D_i} d_k. \tag{4}$$

Since D only contains data held by a unique agent, we have, for each $i \in N$, that

$$b_i^{\min}(N, c_D) = c_D(N) - c_D(N \backslash \{i\}) = -\sum_{k \in D_i} d_k. \tag{5}$$

Subtracting (5) to (4) yields $\sum_{k \in D}(d_k/n)$. Since this expression does not depend on agent $i \in N$, we get $\text{EANSC}_i(N, c_D) - b_i^{\min}(N, c_D) = \text{EANSC}_j(N, c_D) - b_j^{\min}(N, c_D)$ for all $i, j \in N$, as desired. Now, consider any allocation rule f on $\mathcal{DC}$ satisfying the four axioms: to show that f coincides with EANSC. So pick any $(N, c_D) \in \mathcal{DC}$. By Pooling, it is enough to prove that f coincides with EANSC on all elementary games $(N, c_{\{k\}})$, $k \in D$. Assume first that $o_k(N, \{k\}) = 1$, and denote by $i \in N$ the unique owner of data k. It holds that $b_i^{\min}(N, c_{\{k\}}) = c_{\{k\}}(N) - c_{\{k\}}(N \backslash \{i\}) = -d_k$ and that $b_j^{\min}(N, c_{\{k\}}) = c_{\{k\}}(N) - c_{\{k\}}(N \backslash \{j\}) = 0$. Thus, Equal concessions rewrites

$$f_i(N, c_{\{k\}}) + d_k = f_j(N, c_{\{k\}}) \tag{6}$$

for each $j \in N \backslash \{i\}$. Summing over all $j \in N$ and using Efficiency, we obtain

$$0 = \sum_{j \in N} f_j(N, c_{\{k\}}) = f_i(N, c_{\{k\}}) + (n-1)\big(f_i(N, c_{\{k\}}) + d_k\big)$$

so that $f_i(N, c_{\{k\}}) = d_k/n - d_k = \text{EANSC}_i(N, c_{\{k\}})$. By (6), we can conclude that $f_j(N, c_{\{k\}}) = d_k/n = \text{EANSC}_j(N, c_{\{k\}})$ for each $j \in N \backslash \{i\}$ as well. Now consider any $(N, c_{\{k\}})$ with $o_k(N, \{k\}) \geq 2$. Since

[3] Efficiency is renamed Compensation on the class of data games by Béal and Deschamps (2014).

$\text{EANSC}(N, c_{\{k\}}) = \mathbf{0}_n$ in such a case, it remains to show that $f(N, c_{\{k\}}) = \mathbf{0}_n$ whenever $o_k(N, \{k\}) \geq 2$. Note that $o_k(N, \{k\}) \geq 2$ implies that $c_{\{k\}}(N \backslash \{i\}) = 0$ for all $i \in N$. As a consequence, for all $i \in N$, Efficiency in the data subgame $(N \backslash \{i\}, c_{\{k\}|N \backslash \{i\}})$ yields

$$\sum_{k \in N \backslash \{i\}} f_k(N \backslash \{i\}, c_{\{k\}|N \backslash \{i\}}) = c_{\{k\}|N \backslash \{i\}}(N \backslash \{i\}) = c_{\{k\}}(N \backslash \{i\}) = 0.$$

Using this equality and rewriting Balanced collective contributions as in the proof of Proposition 1, we get $f_i(N, c_{\{k\}}) = f_j(N, c_{\{k\}})$ for all $i, j \in N$. Combined with Efficiency in $(N, c_{\{k\}})$, conclude that $f(N, c_{\{k\}}) = \mathbf{0}_n$ as desired. ■

The logical independence of the axioms invoked in Proposition 2, and in turn the fact that Equal concessions and Pooling are both necessary to the result, is demonstrated as follows:

- The allocation rule f on $\mathcal{DC}$ that assigns to each data game $(N, c_D) \in \mathcal{DC}$ the payoff vector $f(N, c_D) = \mathbf{0}_n$ satisfies Efficiency, Balanced collective contributions and Pooling but violates Equal Concessions.
- The Shapley value Sh on $\mathcal{DC}$ satisfies Efficiency, Pooling and Equal Concessions but violates Balanced collective contributions.
- For each data $k \in \mathcal{D}$, choose $a_k \in \mathbb{R}_+$ such that not all a_k are null. The allocation rule f on $\mathcal{DC}$ that assigns to each data game $(N, c_D) \in \mathcal{DC}$ the payoff vector $f(N, c_D) = \sum_{k \in D} a_k + \text{EANSC}(N, c_D)$ satisfies Balanced collective contributions, Pooling and Equal Concessions but violates Efficiency.
- The allocation rule f on $\mathcal{DC}$ that assigns to each data game $(N, c_D) \in \mathcal{DC}$ the payoff vector $f(N, c_D) = \text{Sh}(N, c_D)$ if $\emptyset \subsetneq D^E \subsetneq D$ and $f(N, c_D) = \text{EANSC}(N, c_D)$ otherwise satisfies Efficiency, Balanced collective contributions and Equal Concessions but violates Pooling.

The comparison between the equal allocation of non-separable costs and the Shapley value provided by Proposition 1 and the result in Myerson (1980) can be repeated on the class of data games. In fact, for the same reason as those mentioned above, the characterization of the Shapley value in Myerson (1980) does not hold on the class of data games. Again, Balanced contributions only applies to data games containing no exclusive data, and the null solution satisfies the axiom as well as Efficiency. Nonetheless, replacing Balanced collective contributions by Balanced contributions in Proposition 2 yields a characterization of the Shapley value on the class of data games.

Proposition 3 *The Shapley value is the unique allocation rule on $\mathcal{DC}$ that satisfies Efficiency, Balanced contributions, Pooling and Equal concessions.*

Proof. The Shapley value satisfies Efficiency and Balanced contributions by by Myerson (1980), and Pooling by (3). It also satisfies Equal concessions by the proof of Proposition 2 and the fact that the Shapley value and the equal allocation of non-separable costs coincide on data games in which all data are exclusive. Next let f be an allocation rule on $\mathcal{DC}$ that satisfies the four axioms. Let us show that $f = \text{Sh}$ on $\mathcal{DC}$. By pooling, it is enough to show that $f(N, c_{\{k\}}) = \text{Sh}(N, c_{\{k\}})$ for all $k \in \mathcal{D}$. We proceed by induction of the size of N.
Initial step: if $n = 1$, then $f(N, c_{\{k\}}) = \text{Sh}(N, c_{\{k\}})$ follows from Efficiency.
Induction hypothesis: assume that $f(N, c_{\{k\}}) = \text{Sh}(N, c_{\{k\}})$ for all $k \in \mathcal{D}$ whenever $n < m$ for some finite natural number $m \geq 1$.
Induction step: consider a data game $(N, c_{\{k\}})$, $k \in \mathcal{D}$, such that $n = m$. There are two sub-cases. If $o_k(N, \{k\}) = 1$, from the proof of Proposition 2 and the coincidence of the Shapley value and the equal allocation of non-separable costs, the combination of Efficiency and Equal concessions implies $f(N, c_{\{k\}}) =$

EANSC$(N, c_{\{k\}}) = $ Sh$(N, c_{\{k\}})$. If $o_k(N, \{k\}) \geq 2$, then Balanced contributions can be applied. Together with the induction hypothesis, for any pair of agents $i, j \in N$, it holds that

$$\begin{aligned}\mathrm{Sh}_i(N, c_{\{k\}}) - \mathrm{Sh}_j(N, c_{\{k\}}) &= \mathrm{Sh}_i(N\backslash\{j\}, c_{\{k\}|N\backslash\{j\}}) - \mathrm{Sh}_j(N\backslash\{i\}, c_{\{k\}|N\backslash\{i\}}) \\ &= f_i(N\backslash\{j\}, c_{\{k\}|N\backslash\{j\}}) - f_j(N\backslash\{i\}, c_{\{k\}|N\backslash\{i\}}) \\ &= f_i(N, c_{\{k\}}) - f_j(N, c_{\{k\}}).\end{aligned}$$

Since the equality holds for all pairs $i, j \in N$, there must exist a constant $a \in \mathbb{R}$ such that $\mathrm{Sh}_i(N, c_{\{k\}}) - f_i(N, c_{\{k\}})$ for all $i \in N$. By Efficiency of both Sh and f, we obtain $a = 0$ as desired, and the proof is complete. ∎

The four the axioms invoked in Proposition 3 are also logically independent as shown by the following allocation rules:

- The allocation rule f on $\mathcal{DC}$ that assigns to each data game $(N, c_D) \in \mathcal{DC}$ the payoff vector $f(N, c_D) = \mathbf{0}_n$ satisfies Efficiency, Balanced contributions and Pooling but violates Equal Concessions.
- The equal allocation of non-separable costs on $\mathcal{DC}$ satisfies Efficiency, Pooling and Equal Concessions but violates Balanced contributions.
- For each data $k \in \mathcal{D}$, choose $a_k \in \mathbb{R}_+$ such that not all a_k are null. The allocation rule f on $\mathcal{DC}$ that assigns to each data game $(N, c_D) \in \mathcal{DC}$ the payoff vector $f(N, c_D) = \sum_{k \in D} a_k + \mathrm{Sh}(N, c_D)$ satisfies Balanced contributions, Pooling and Equal Concessions but violates Efficiency.
- The allocation rule f on $\mathcal{DC}$ that assigns to each data game $(N, c_D) \in \mathcal{DC}$ the payoff vector $f(N, c_D) = $ EANSC(N, c_D) if $\emptyset \subsetneq D^E \subsetneq D$ and $f(N, c_D) = \mathrm{Sh}(N, c_D)$ otherwise satisfies Efficiency, Balanced contributions and Equal Concessions but violates Pooling.

5 Conclusion

We conclude this article by suggesting another possible development close to our study. Since the class of data games is a subclass of the class of 1-concave games, it would be interesting to check whether the results in section 4 still hold on this larger class. In particular, on this larger class, is it necessary to add extra axioms to those invoked in Proposition 1 and the result in Myerson (1980) as we did in section 4 to recover the characterizations of the equal allocation of non-separable costs and the Shapley value, respectively?

References

Béal, S., Deschamps, M., 2014. On compensation schemes for data sharing within the european REACH legislation, CRESE working paper no 2014-01.

Brânzei, R., Tijs, S. H., Timmer, J., 2006. Compensations in information collecting situations: A cooperative approach. Journal of Public Economic Theory 8, 181–191.

Dehez, P., Tellone, D., 2013. Data games: Sharing public goods with exclusion. Journal of Public Economic Theory 15, 654–673.

Driessen, T. S. H., 1988. Cooperative Games, Solutions, and Applications. Kluwer Academic, Dordrecht, The Netherlands.

Driessen, T. S. H., Khmelnitskaya, A., 2013. A comment on P. Dehez and D. Tellone "Data games: sharing public goods with exclusion", memorandum 2013, Department of applied mathematics, University of Twente, Enschede, The Netherlands.

Driessen, T. S. H., Khmelnitskaya, A. B., Sales, J., 2012. 1-concave basis for TU games and the library game. TOP 20, 578–591.

Heaney, J. P., 1979. Efficiency/equity analysis of environmental problems: A game theoretic perspective. In: Applied Game Theory. S. J. Brams, A. Schotter and G. Schwodiauer (Eds.), Physica-Verlag, Vienna, pp. 352–369.

Herings, P. J.-J., van der Laan, G., Talman, A. J. J., 2008. The average tree solution for cycle-free graph games. Games and Economic Behavior 62, 77–92.

Myerson, R. B., 1980. Conference structures and fair allocation rules. International Journal of Game Theory 9, 169–182.

Ransmeier, J. S., 1942. The tennessee valley authority: A case study in the economics of multiple purpose stream planning, the Vanderbuilt University Press, Nashville, Tennessee.

Schmeidler, D., 1969. The Nucleolus of a characteristic function game. SIAM Journal of Applied Mathematics 17, 1163–1170.

Shapley, L. S., 1953. A value for n-person games. In: Contribution to the Theory of Games vol. II (H.W. Kuhn and A.W. Tucker eds). Annals of Mathematics Studies 28. Princeton University Press, Princeton.

Straffin, P. D., Heaney, J. P., 1981. Game theory and the tennessee valley authority. International Journal of Game Theory 10, 35–43.

Thomson, W., 2001. On the axiomatic method and its recent applications to game theory and resource allocation. Social Choice and Welfare 18, 327–387.